Machine Learning Tools for Predicting Solar Energetic Particle Hazards: Progress and Plans

Viacheslav Sadykov

NASA Ames Research Center

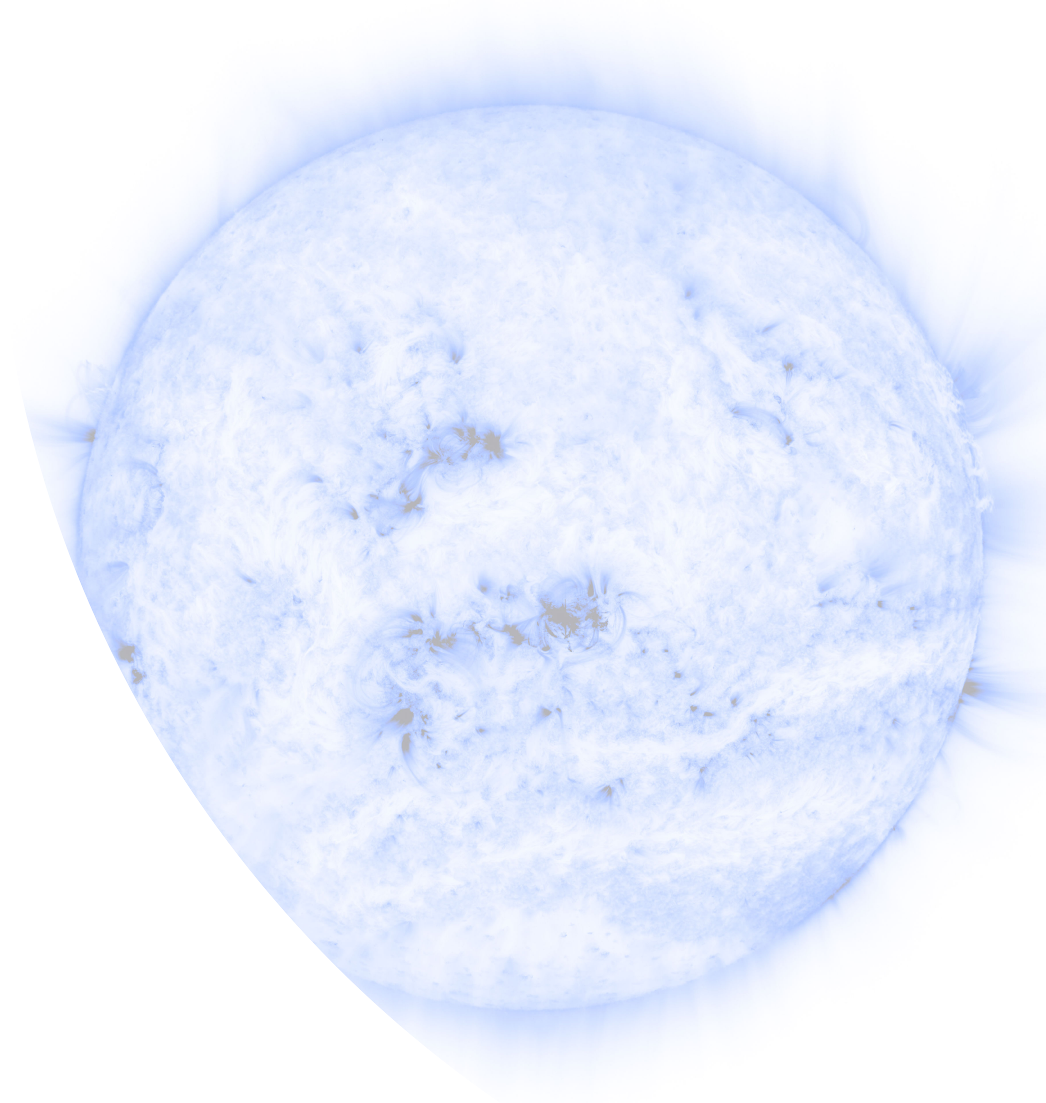
Bay Area Environmental Research Institute

AMS Seminar Series, NASA ARC

November 5, 2020

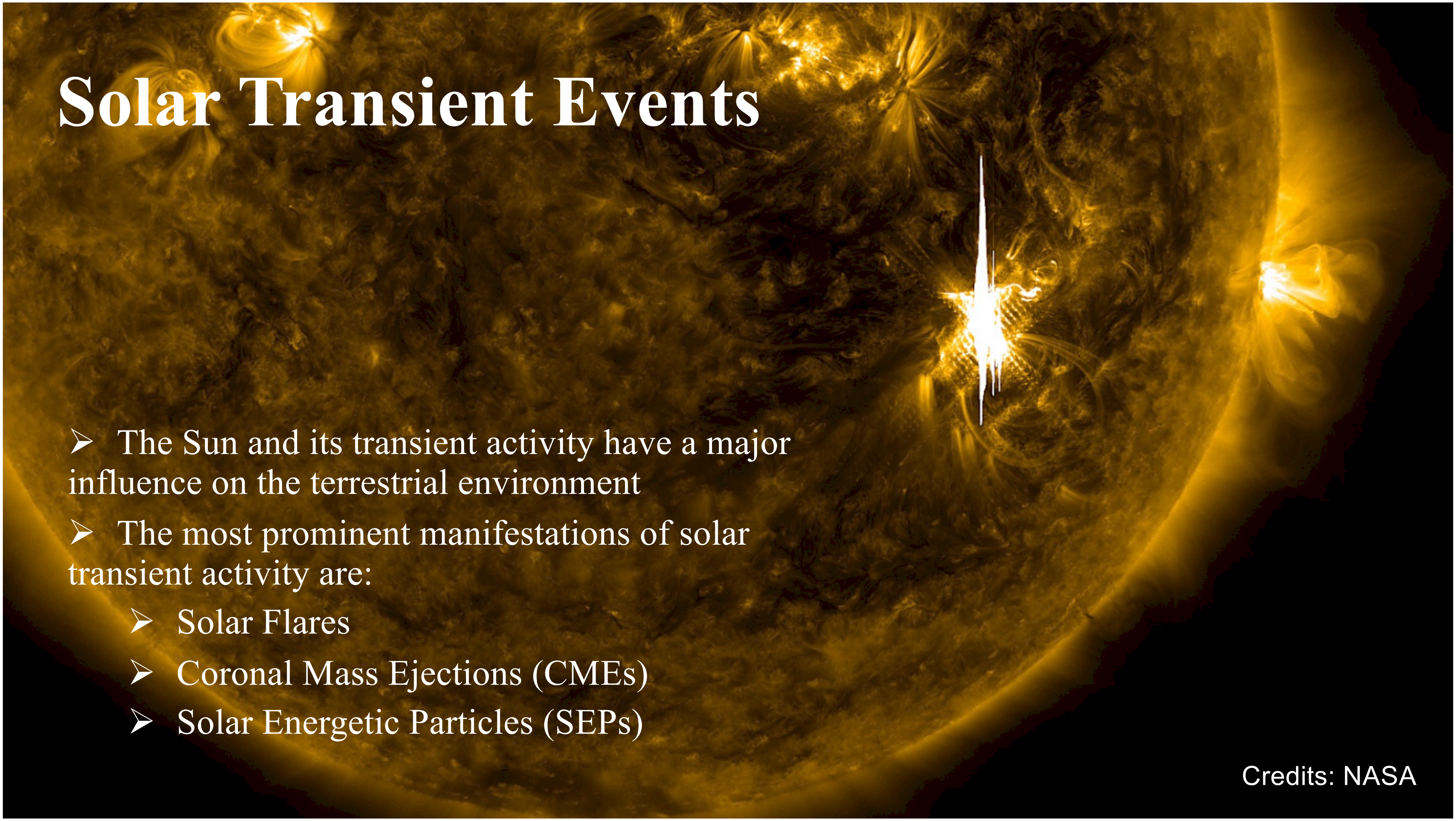
Outline

- Introduction: prediction of Solar Energetic Particle (SEP) events
- ESI project: goals and objectives of the proposed research
- Part 1. An online-accessible database of SEP-related solar and heliospheric data, metadata, and descriptors
- Part 2. Development of “all-clear” forecasts of Solar Proton Events (SPEs)
- Conclusions



Solar Transient Events

- The Sun and its transient activity have a major influence on the terrestrial environment
- The most prominent manifestations of solar transient activity are:
    - Solar Flares
    - Coronal Mass Ejections (CMEs)
    - Solar Energetic Particles (SEPs)

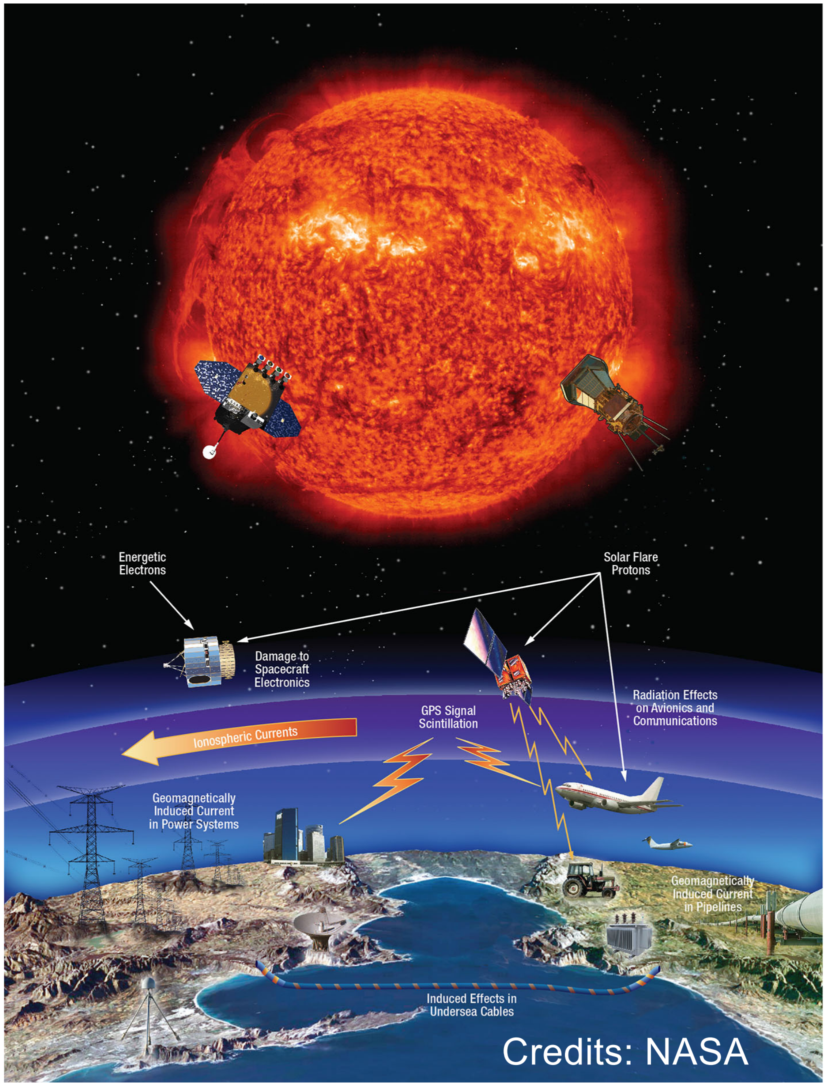


Credits: NASA

Space Weather impact of Solar Energetic Particles

- SEPs can damage spacecraft electronics
- Radiation exposure at aviation altitudes increases during SEP events
- SEPs represent a potential danger for space exploration

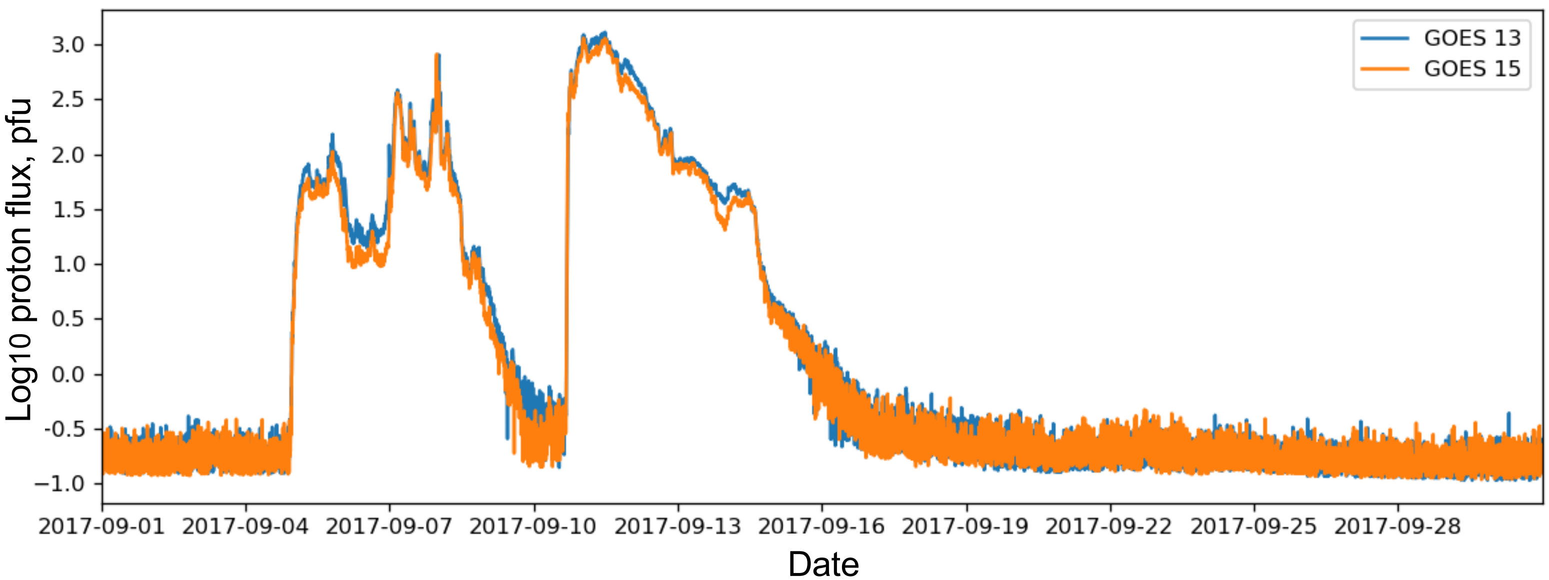
Prediction of SEPs is a critically important problem from both operational and research perspectives



Solar Energetic Particles (SEPs) and Solar Proton Events (SPEs)

- Solar Energetic Particle (SEP) events can be defined as significant enhancements of the particle flux coming from the Sun with respect to the stable background
- Solar Proton Events (SPEs) represent a major subclass of SEPs
- The terms "SEP event" and "SPE" are equivalent for this presentation and represent enhancements of energetic proton fluxes as measured by near-Earth satellites (GOES)

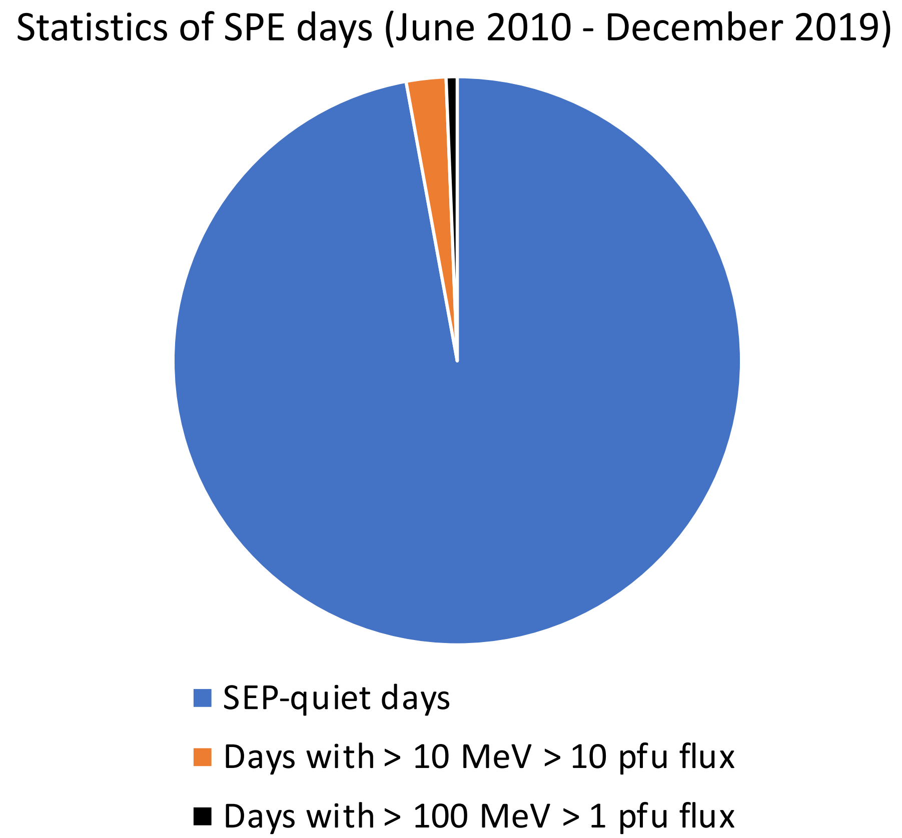
> 10 MeV proton flux



An example of > 10 MeV proton flux measurements by the GOES-13 and GOES-15 satellites

Why is predicting solar proton events challenging?

- Severe class-imbalance ratio. The ratio of SPE-active to SPE-quiet days is:
  - **1/34** for > 10 MeV > 10 pfu events
  - **1/155** for > 100 MeV > 1 pfu events
- SPE onset may occur significantly later than the initiating flare.
- The locations of SPE initiations on the Sun are not known precisely. Some events are initiated on the far side of the solar disk.



Encouragement of collaborative effort

- NASA encourages the development of SEP prediction capabilities:
  - NASA Space Weather Operations to Research (SWO2R) program elements
  - NASA Early Stage Innovation (ESI) program elements
  - Other broadly-defined NASA solicitations and programs
- Discussions between various SEP prediction groups has been established in the US and elsewhere, for both research and operations.
- Interested in joining the ML SEP group monthly discussions? Contact Dr. Irina Kitiashvili (irina.n.kitiashvili@nasa.gov, meeting organizer) to be added to the mailing list.

ESI Project: Machine Learning Tools for Predicting Solar Energetic Particle Hazards

PI: Alexander Kosivichev (NJIT)

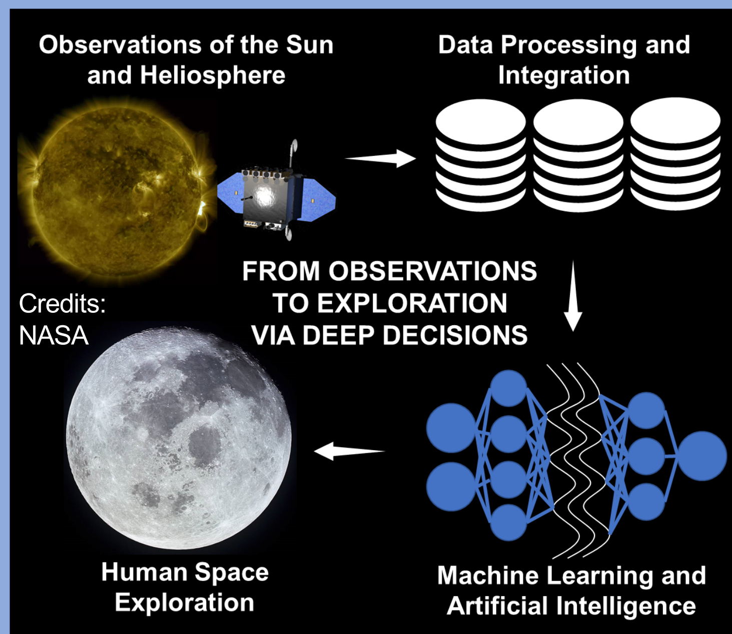
Co-Is: Vincent Oria (NJIT), Viacheslav Sadykov (BAERI), Irina Kitiashvili (NASA ARC)

Students (NJIT): Yucheng Jiang, Patrick O'Keefe, Sheldon Fereira

Collaborator: Egor Illarionov (MSU)

Research Objectives

- Enhance predictions of solar energetic particles (SEP) by developing an online-accessible database of SEP-related data and implementing robust machine-learning-based "all-clear" forecasts;
- Innovation: unleashing machine learning technologies for SPE forecasts, enabling data discovery
- Enhancement of SOA: automatic forecasts without "forecaster-in-the-loop"



Approach

- Modern Machine Learning and database technologies with API-based online access to database entries and integrated data products
- Automated utilization and processing of multi-spacecraft observational data
- Customized Skill Score for "all-clear" forecasts
- Application of advanced Machine Learning algorithms for solving classification tasks and prediction of active region evolution

Potential Impact

- Robust "all-clear" forecasts, enhancing space exploration safety and advanced mission planning
- Availability of prepared and integrated data for the heliophysics community
- Data discovery possibilities for the broader community
- Replacement of current operational forecasting techniques by machine-learning-based approaches

Goals and Objective of the proposed research

The primary objective of the proposed research is to enhance predictions of solar energetic particles (SEP) by implementing automatic data characterization and machine-learning tools.

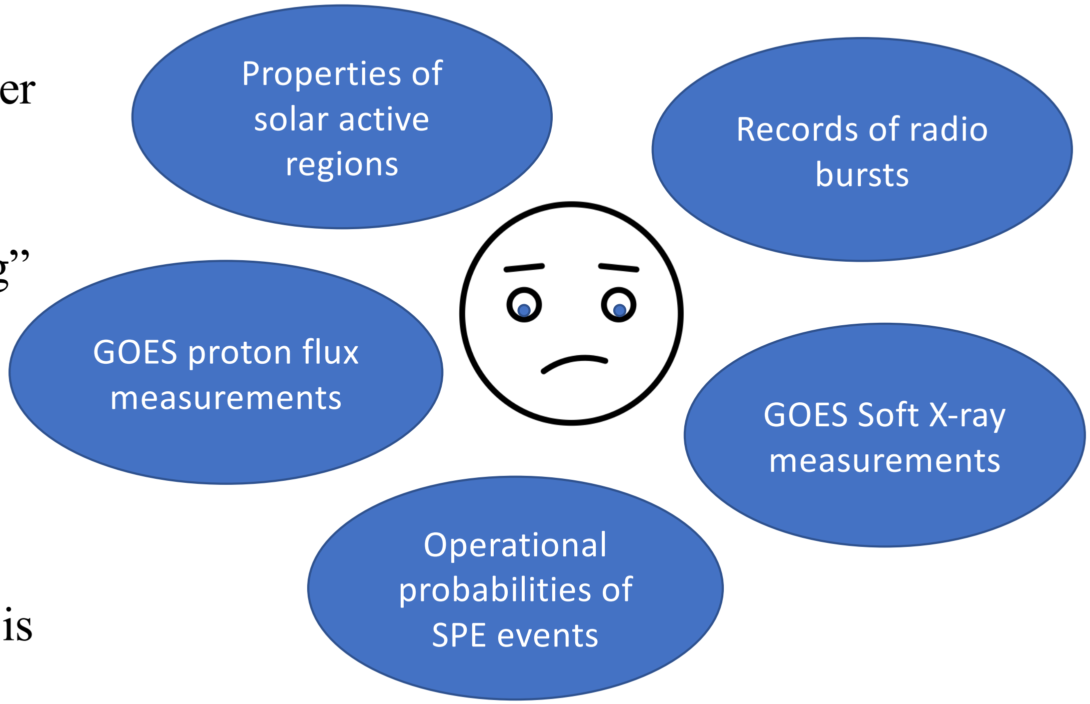
The proposal pursues two main goals:

- Develop an online-accessible, automatically-updated database that integrates the solar and heliospheric data, metadata, and descriptors related to SPEs.
- Develop robust "all-clear" forecasts of SPEs with low false-alarm rates, targeted at different temporal scales (cadences and lead times), different energy and particle flux thresholds of SPEs, and adapted to the operational availability of data sources and gaps in the data.

Online-accessible database for SPE-related data

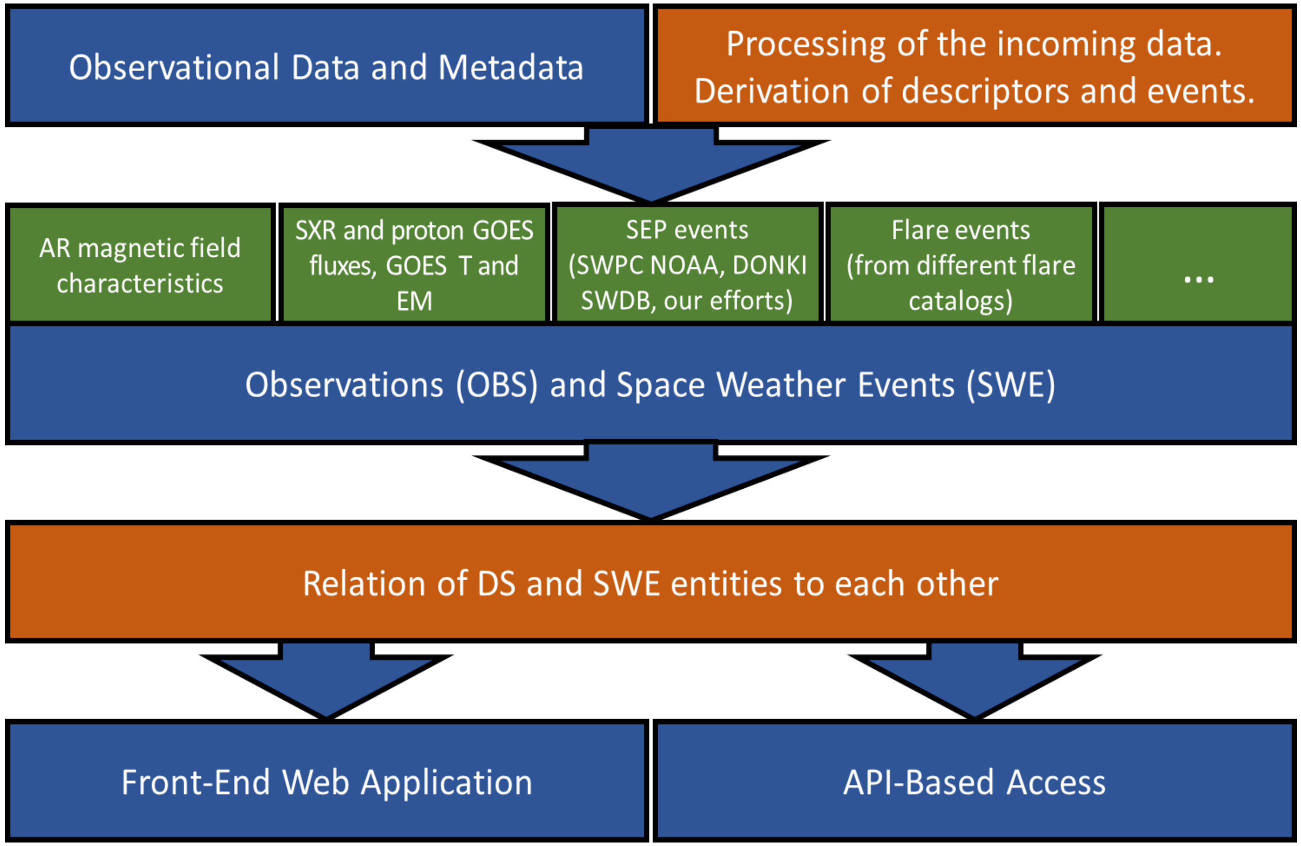
On the situation with SPE-related data

- The data and event records are often stored in individual catalogs spread over various locations (excepting some efforts).
- These catalogs are "not always talking" to each other.
- The data sources may contain errors and therefore must be cross-checked.
- Any prediction attempt (including machine learning) starts from data collection and preparation. This phase is very time-consuming.



Our approach to data management

- We are currently developing an online-accessible, automatically updated database of SPE-related solar and heliospheric data, metadata, and descriptors.
- The idea is to bring the SPE-related data together and provide easy access and data integration.
- Most of the database work is done by graduate students in the Physics and Computer Science departments at NJIT

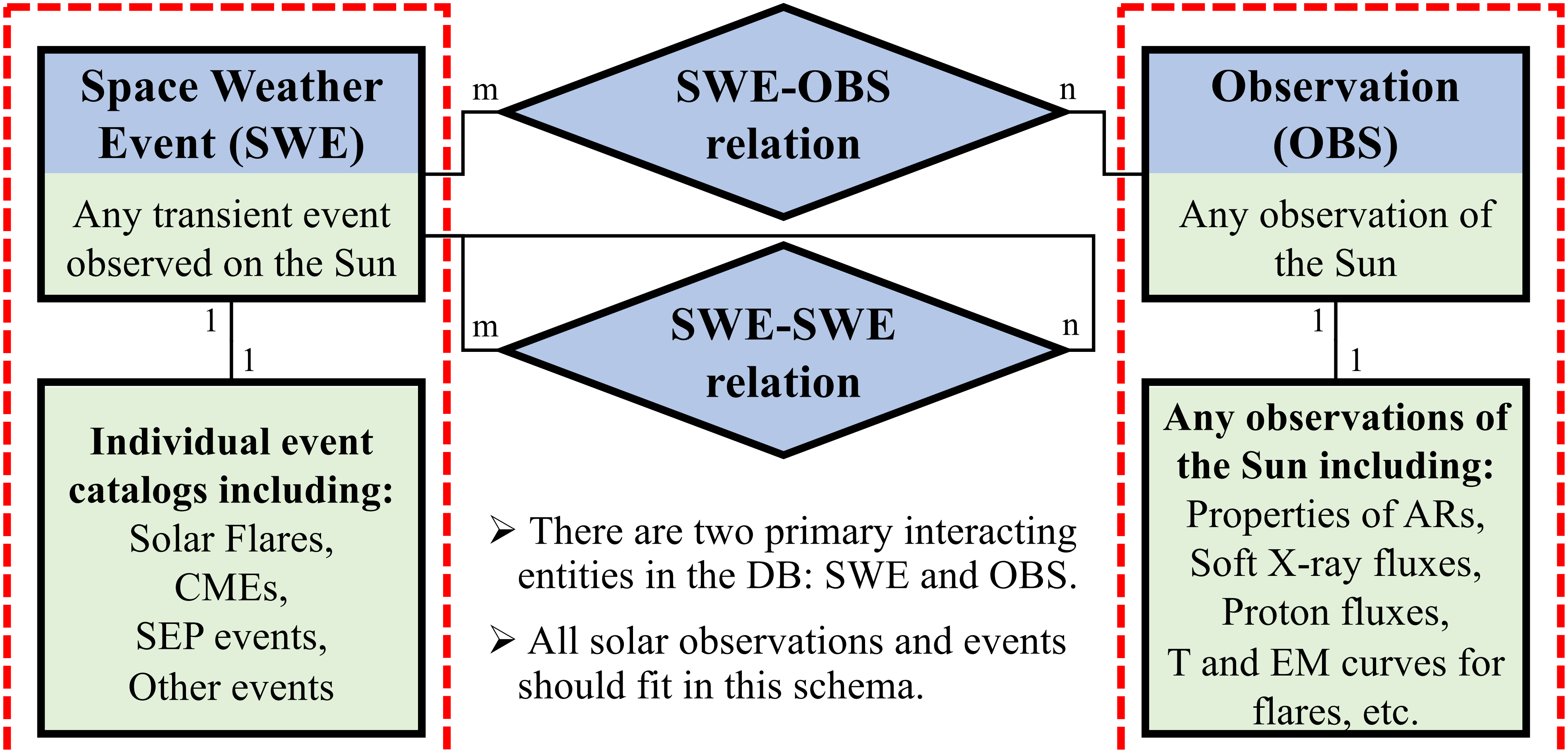


Schematic representation of the database for machine-learning prediction of Solar Proton Events (SPE).

Collected data sources

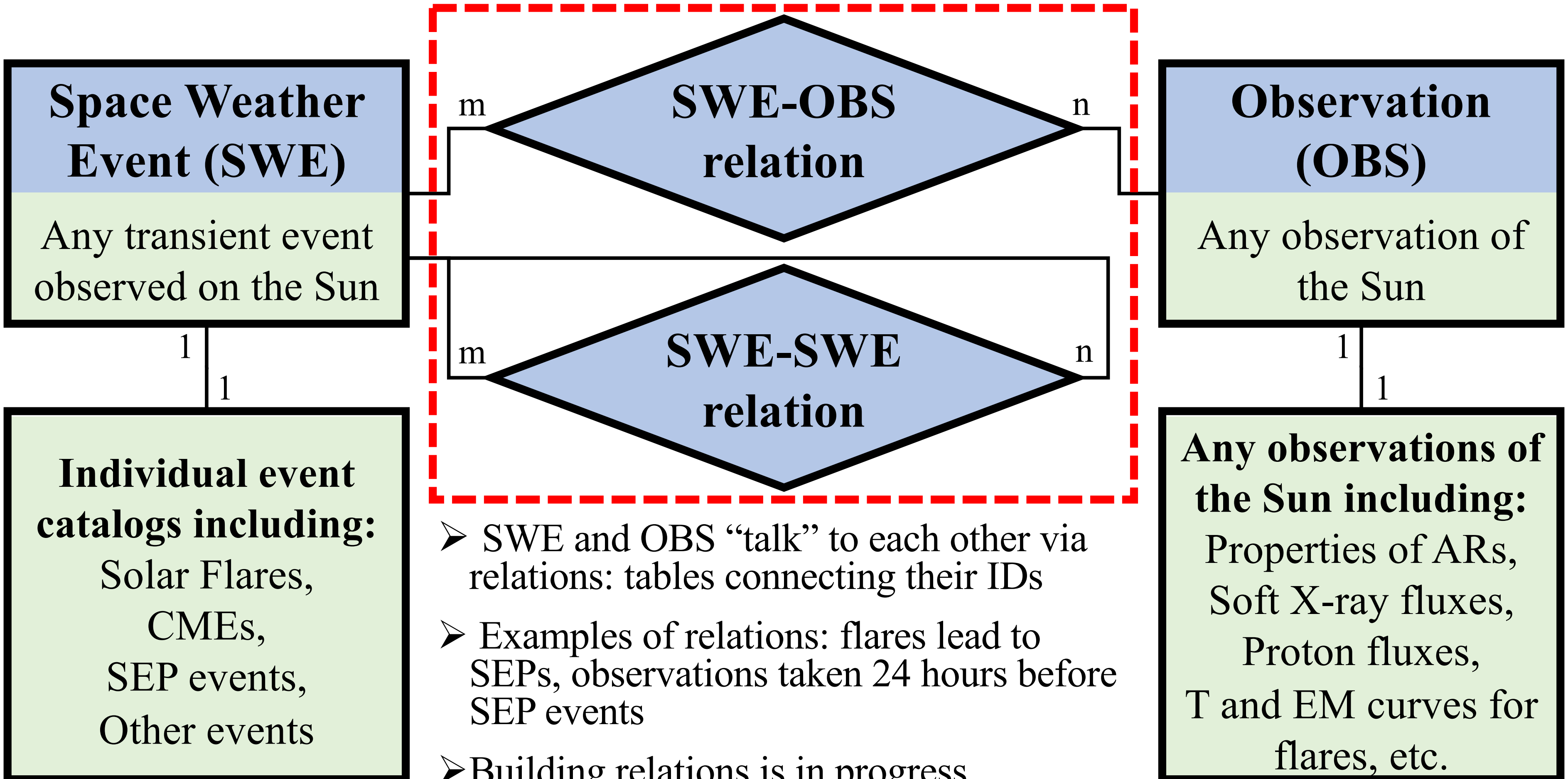
Solar Energetic Particle DB Data Sources (2010-2020)	
Polarity Inversion Line (PIL) properties of Active Regions	GOES proton flux measurements
Space Weather HMI Active Region Patches (SHARPs)	GOES soft X-ray measurements
Flare lists from various flare catalogs (GOES, RHESSI, HEK)	GOES Temperatures and Emission measures
CME lists (LASCO/SOHO and CACTus catalogs)	Records of radio bursts (type II and IV)
Measurements of ground-based neutron monitors	Derived lists of SPE events and SPE-quiet periods

Simplified Entity Relationship Diagram



- There are two primary interacting entities in the DB: SWE and OBS.
- All solar observations and events should fit in this schema.

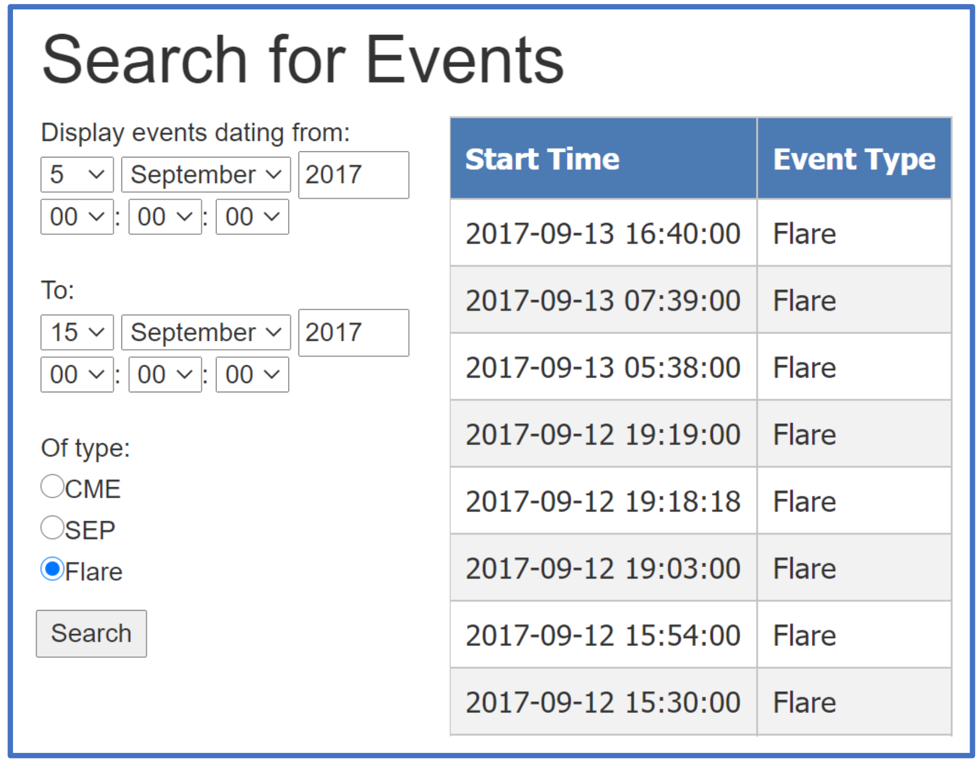
Simplified Entity Relationship Diagram



- SWE and OBS "talk" to each other via relations: tables connecting their IDs
- Examples of relations: flares lead to SEPs, observations taken 24 hours before SEP events
- Building relations is in progress.

API and Front-End Development

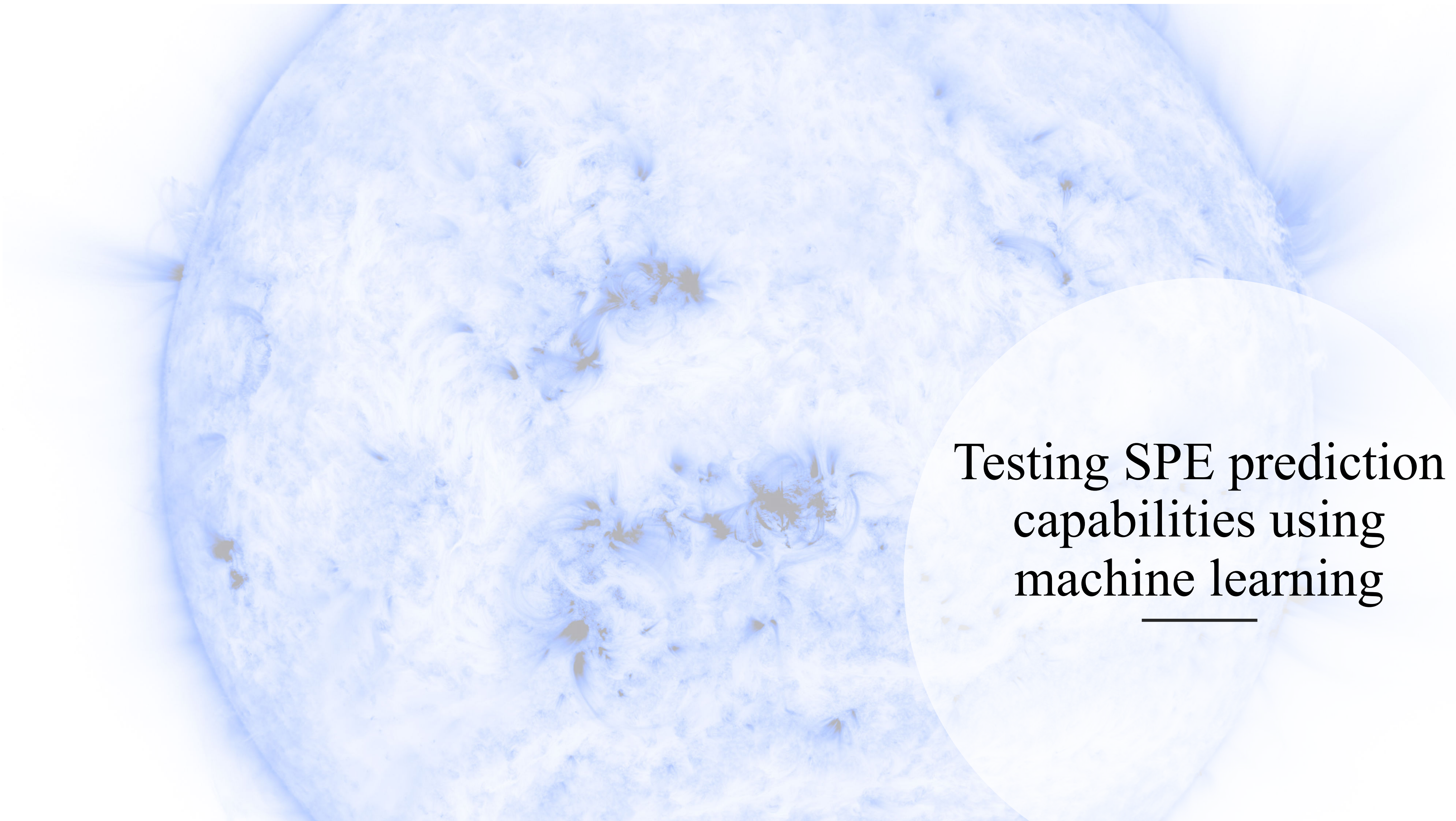
- The web application provides a graphical interface to query our database and get SEP-related data.
- An Application Programming Interface (API) is used to interact with the DB. The DB can be queried directly using the API.
- Planned web app capabilities:
  - Search for events based on their properties (in progress)
  - Search for events related to a particular event (in progress)
  - Search and retrieval of event-related observational data
  - Quicklook visualizations for events



A demonstration of the web application

Database development plans

- Include more SEP-relevant data sources
- Construct more relations between Space Weather Events and Observations for:
  - Predictions of SPEs of different energy and particle flux thresholds at different timescales
  - Different relational models (time-based, location-based, etc.)
- Make the database automatically updated on a daily basis
- Continue developing a functional web interface to query and visualize the data



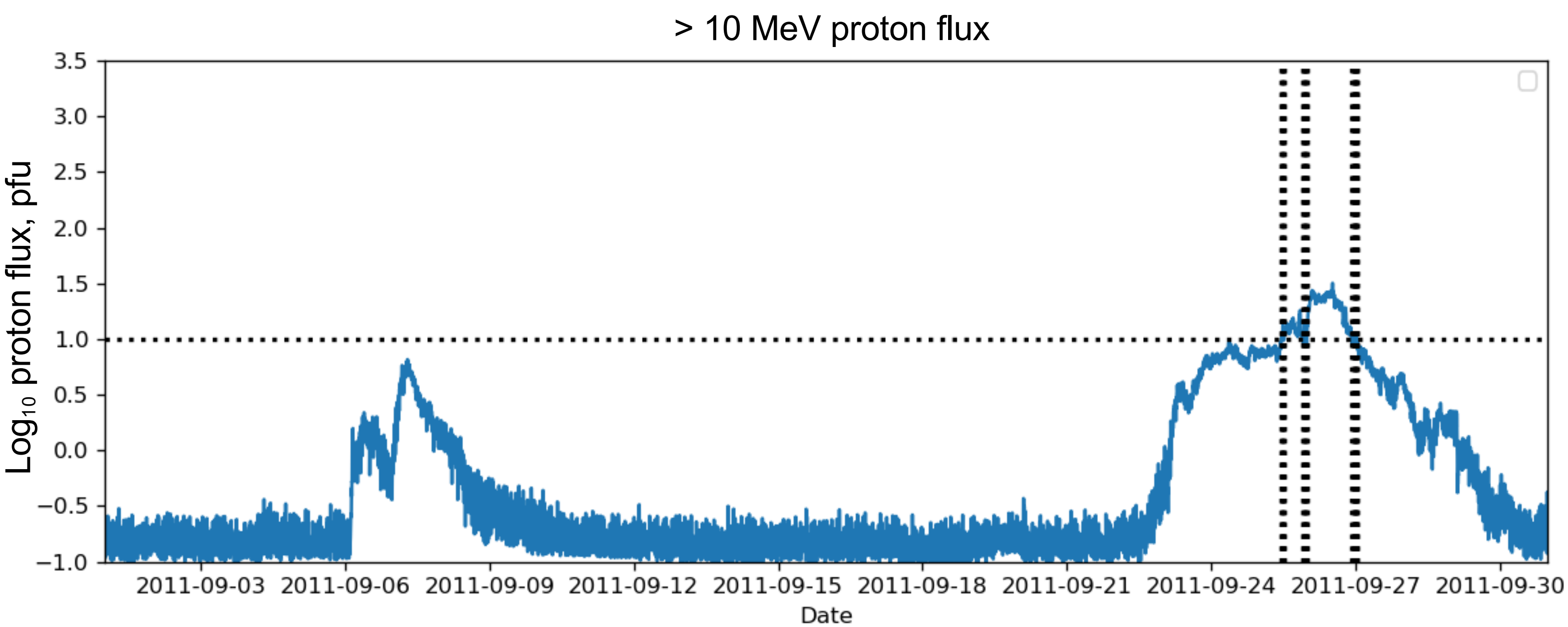
Testing SPE prediction capabilities using machine learning

What does it mean to predict an SPE event?

- In the framework of this study, to predict an SPE event means, for example:

- To predict at 12 AM UT whether the measured peak flux of > 10 MeV protons will exceed 10 particle flux units during the next day. Timeframe: June 2010 – December 2019.

- Defined in this way, the predictions can be compared directly with the SWPC NOAA operational daily forecasts.



An example of > 10 MeV proton flux measurements by the GOES-15 satellite

Machine learning perspective on the problem

- The problem is a classic binary classification problem.
- To solve this problem, we employ neural networks and minimize the cross-entropy loss function during training. Given the true label y (1 or 0) and the predicted probability of the event p:

$$Loss = -(y \cdot log(p) + (1 - y) \cdot \log(1 - p))$$

- The binary classification results may be represented as a confusion table:

Confusion Matrix	Prediction: SPE event	Prediction: no SPE event
Reality: SPE event	True Positives (TP)	False Negatives (FN)
Reality: no SPE event	False Positives (FP)	True Negatives (TN)

- The binary outcomes can be combined to form metrics like True Skill Statistics (TSS)

$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}$$

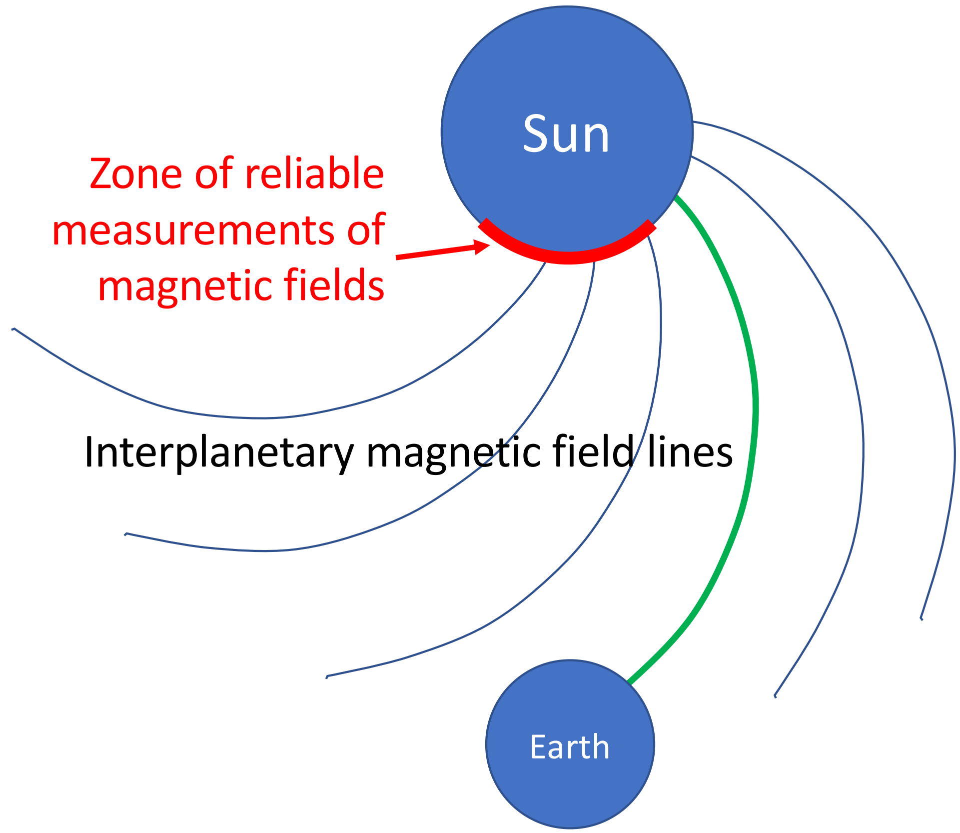
More about the SWPC NOAA operational SPE forecasts

- Issued at 22:00 PM UT for the next day
- Mainly statistics-based (utilize lookup tables and event prehistory)
- Data utilized for SPE forecast: integrated SXR flux, AR locations, presence of type-II and type-IV radio bursts
- 1% is the smallest probability level issued.
- The calculated probabilities can be corrected by forecasters based on their experience.
- The daily forecasts for the whole Sun are available online

- Major problem: during 2010-2020, 14 out of 101 SPE days happened when a 1% chance of the event was predicted. It is problematic to build all-clear forecasts based on that data.

Working with AR information

- The energy released during transient events is (in most cases) initially stored in non-potential configurations of magnetic fields in active regions (ARs).
- SHARP features represent the properties of the vector magnetic field extracted for AR patches (Bobra et al. 2014).
- We utilize the last reliable daily median values of the SHARP AR parameters and assume the AR to have these parameters while traveling behind the limb.



Extracted features

- Median values of SHARP properties for 10 ARs with the largest unsigned magnetic fluxes present on the Sun (including ARs behind the limb)
- Daily properties of SEP flux (mean, median, min, max, and last values, calculated for >10 MeV flux only)
- Daily properties of SXR flux (mean, median, min, max values, for fluxes in both the 0.5 – 4 Å and 1 – 8 Å channels)
- Statistics of Radio Bursts (number of type-II and type-IV bursts)

- Comparison with: SWPC NOAA daily operational forecasts

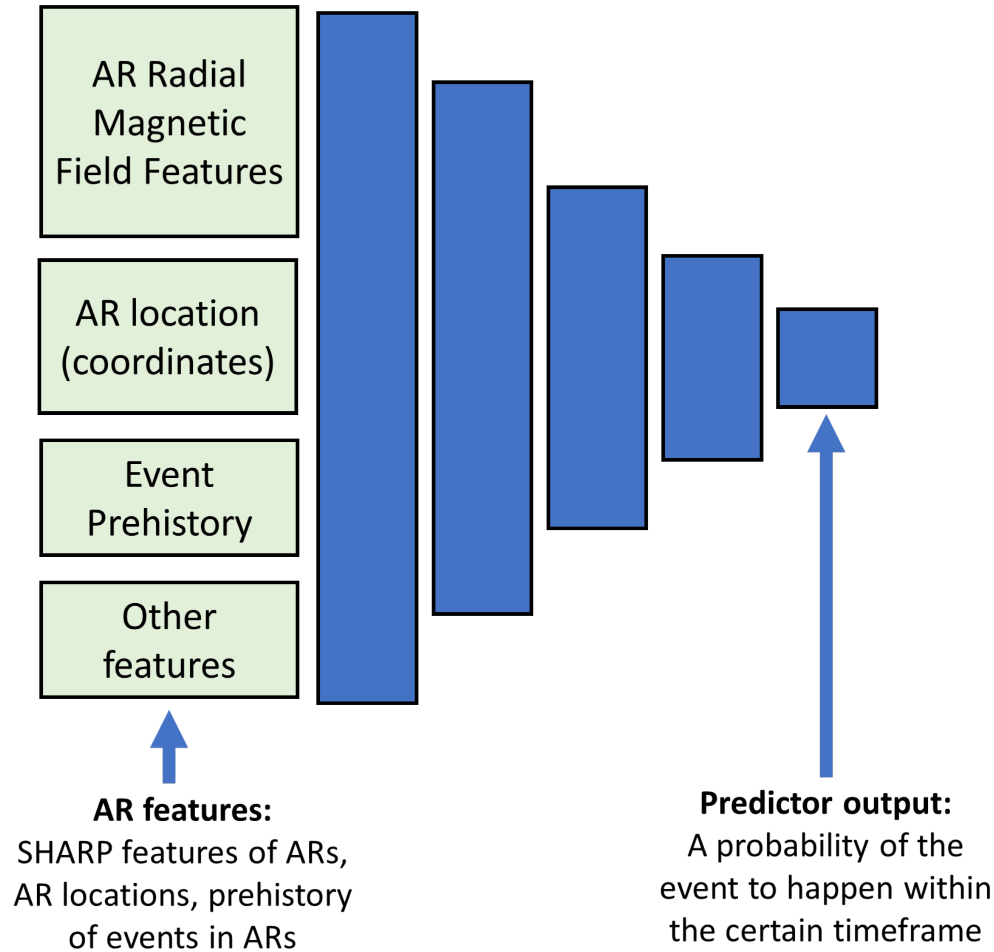
“Classical” prediction for individual ARs

- The probability of an SPE event is issued for each AR present on the solar disk.

Problems:

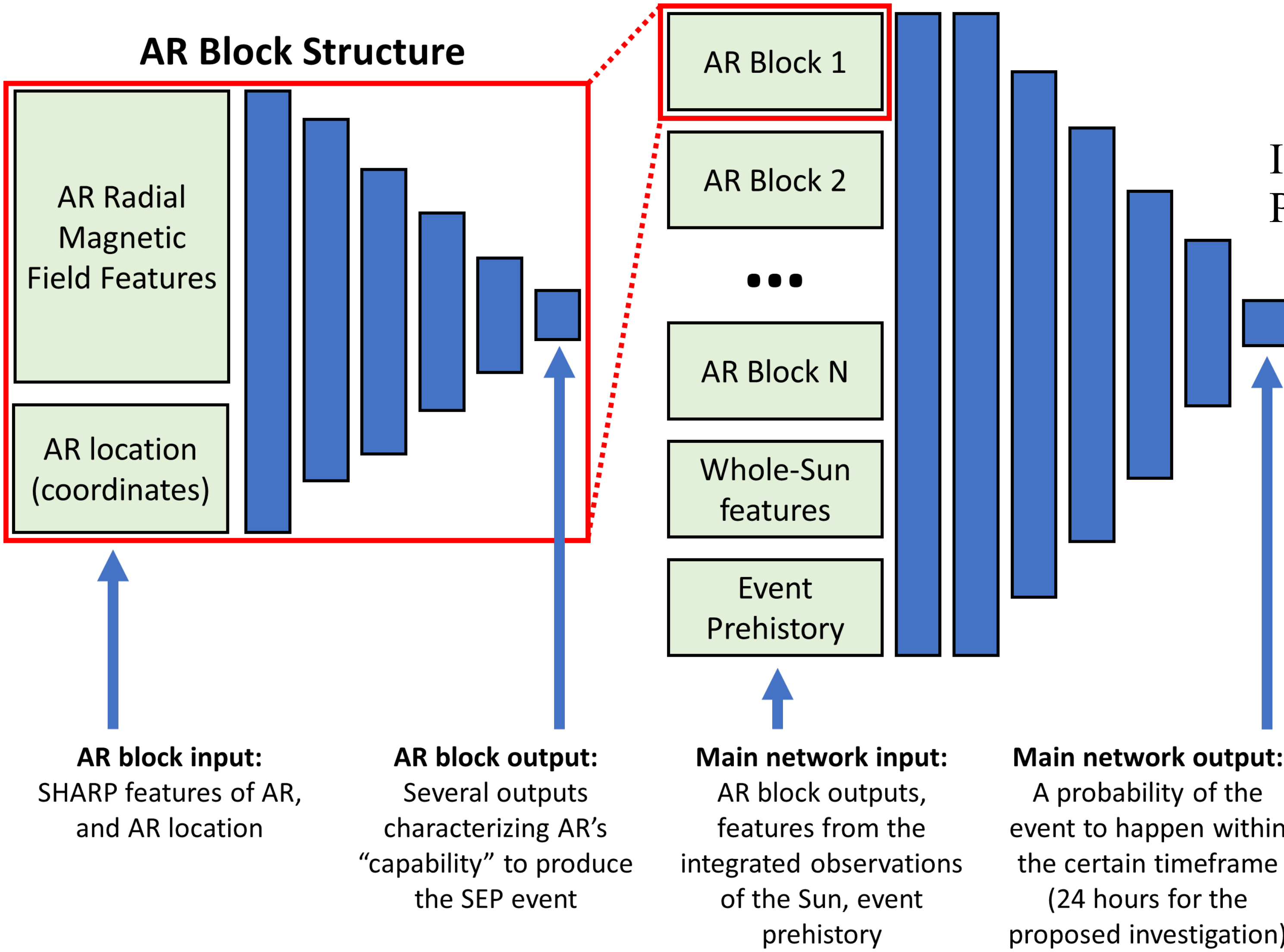
- The class-imbalance ratio is higher for per-AR prediction with respect to the whole-Sun prediction.
- The origin AR is not known for some SEP events.
- Problematic to compare with the whole-Sun forecasts

Prediction structure for individual ARs



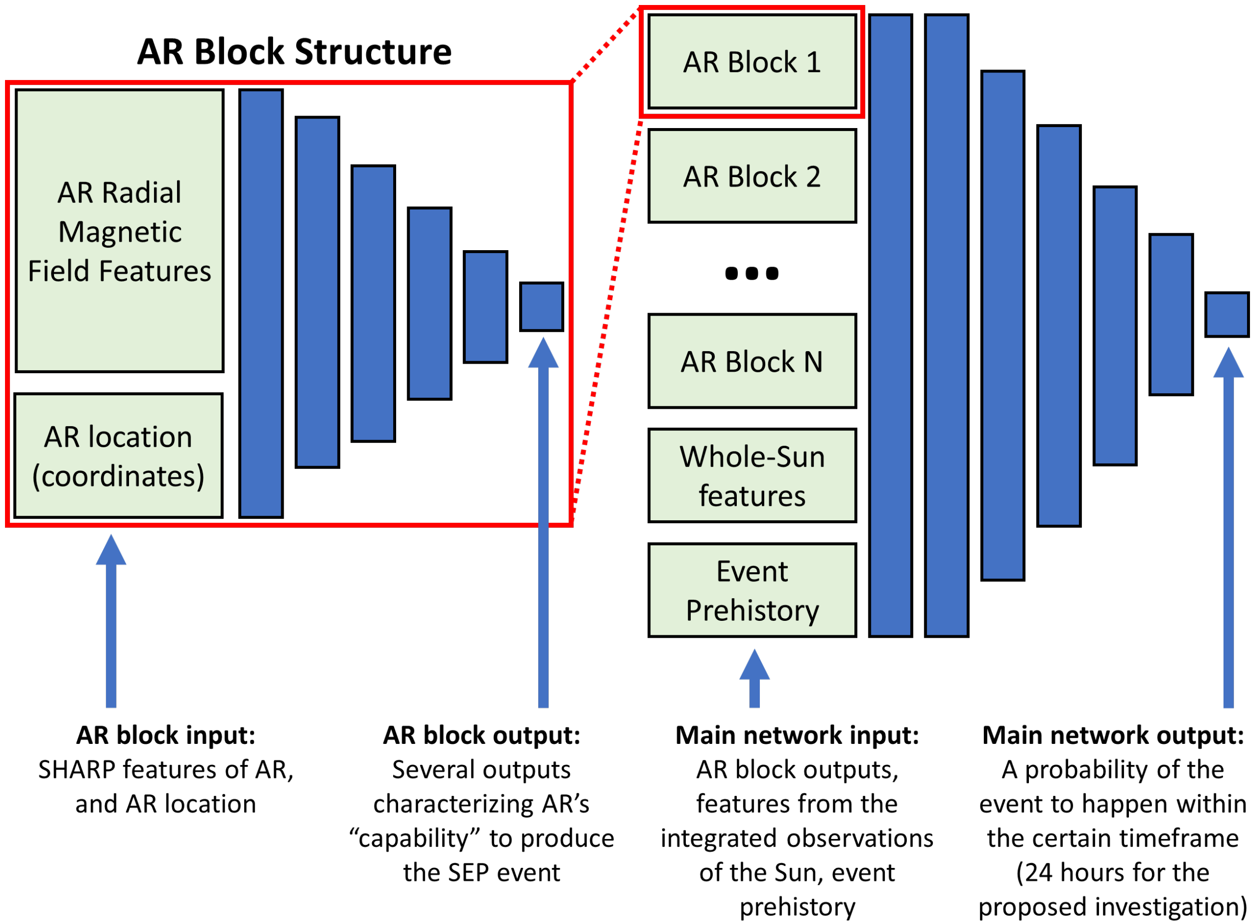
Solution

Implemented using Python PyTorch



Neural network architecture for whole-Sun SEP prediction

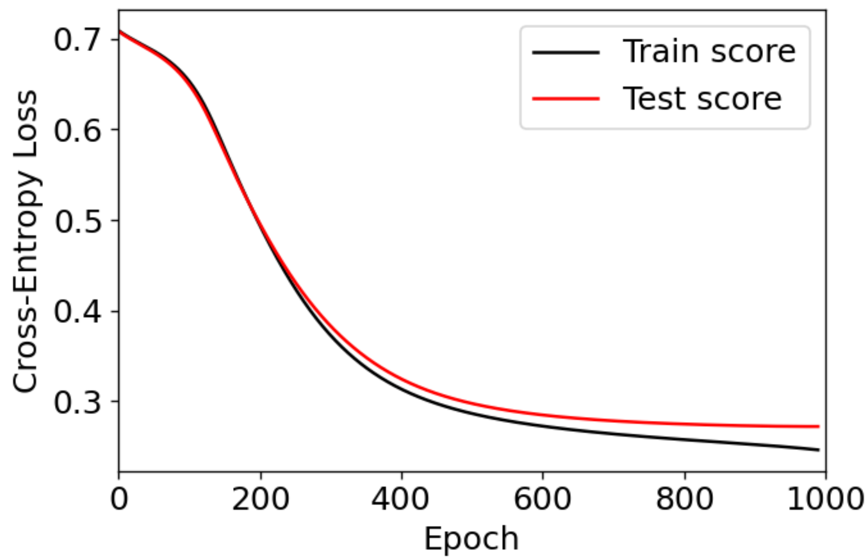
- AR features are processed in "AR Blocks". The weights are shared between the blocks.
- The number of AR Blocks remains the same for each day. The ARs with the highest magnetic fluxes serve as input.
- Whole-Sun features do not need to be linked to the ARs.
- The presented architecture allows us to address the problem of undefined-origin ARs for some SEP events.



Train-test separation and learning strategy

- Time periods in the training data set: 2010-2013, 2016, end of 2018-2019 (66 SPE days)
- Test data set: 2014-2015, 2017-beginning of 2018 (35 SPE days)
- An early stopping criterion is implemented on the test data set to prevent overfitting.
- The developed architecture is much more stable with respect to the fully-connected implementation.
- The procedure was performed 5 times for each investigated setup.

Important note: our goal is not to evaluate our predictor on the unknown data but to investigate how much we can learn from the available data in principle.

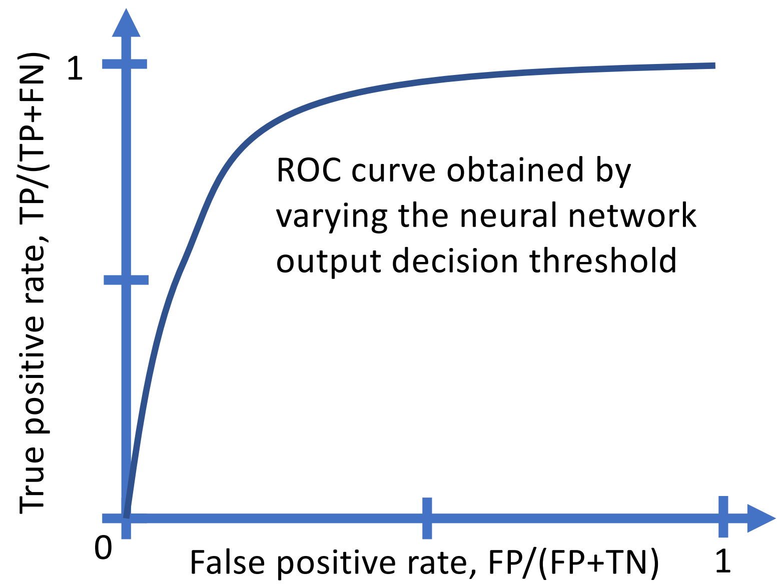


An example of cross-entropy loss for train and test data sets during the training progress

Inclusion/exclusion of parameters during the testing phase

- We would like to investigate how inclusion/exclusion of various parameters affects the prediction:
  - Instead of adapting the network architecture to variable input, we “erase” the information for excluded descriptors (i.e., set the corresponding input to a constant unchanging value)

- Questions to be investigated:
  - Comparison of the neural network prediction with SWPC NOAA forecast
  - Exploration of the prediction solely based on SHARP properties
  - Understanding the role of SHARP and proton flux properties in the prediction
  - Exploration of the Receiver Operating Characteristic (ROC) curves



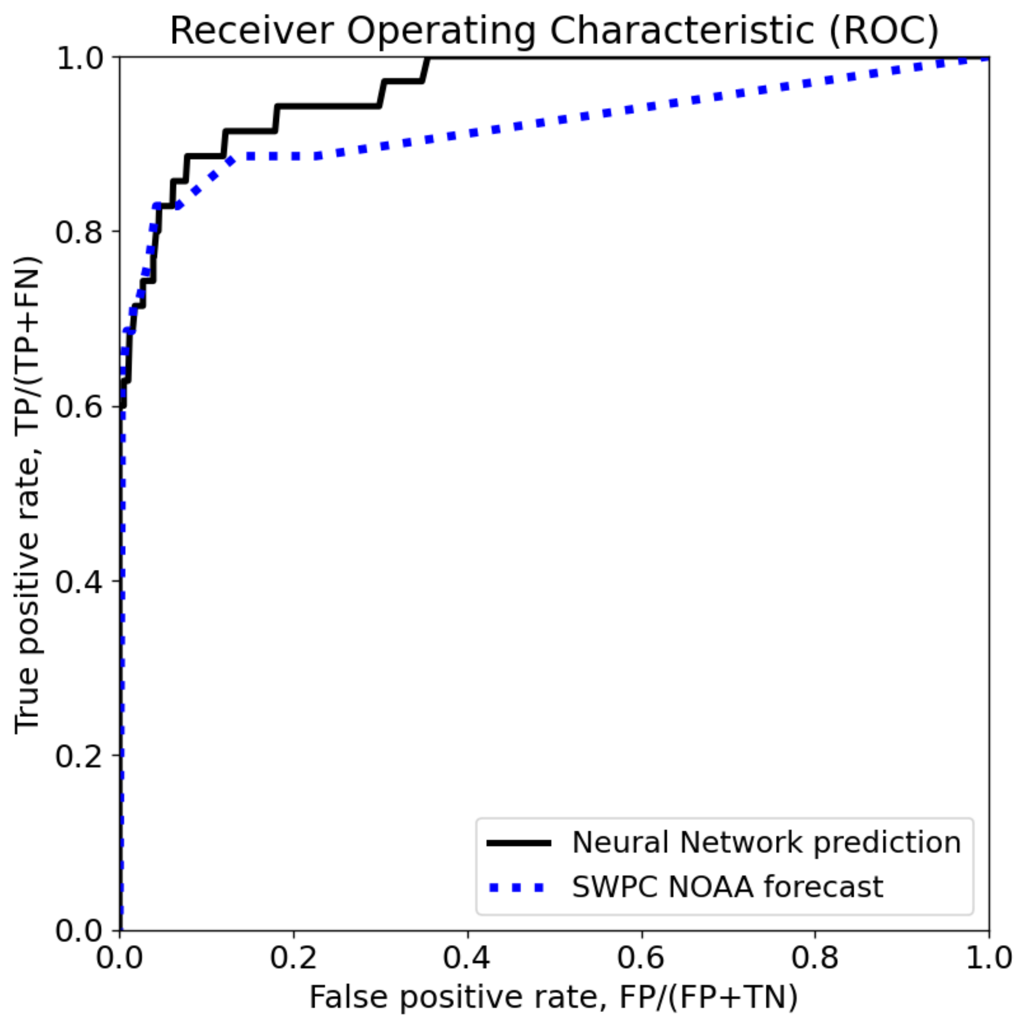
Comparison of ML prediction with SWPC NOAA and persistence forecasts

	Cross-entropy loss	TSS	Best testing example			
			TP	TN	FP	FN
ML prediction	0.271±0.005	0.775±0.004	31	1054	124	4
SWPC NOAA forecast	0.317	0.772	29	1111	67	6
Persistence model	-	0.647	23	1166	12	12

- Both the ML prediction and SWPC NOAA forecast are better than the persistence model (in terms of TSS).
- The ML prediction has the same TSS score as the SWPC NOAA forecast but has a lower cross-entropy loss.
- Let us now look at Receiver Operating Characteristic (ROC) curves for the network

ROC curves for the forecasts

- Although the TSS scores of both forecasts were the same, the ROC curves show a difference.
- The ML-based forecast clearly outperforms the SWPC NOAA forecast at the higher true positive rates relevant to all-clear forecasts.
- 4 SPE events were totally missed by SWPC NOAA forecasts: the corresponding issued probabilities were 1% (lowest-issued probability).



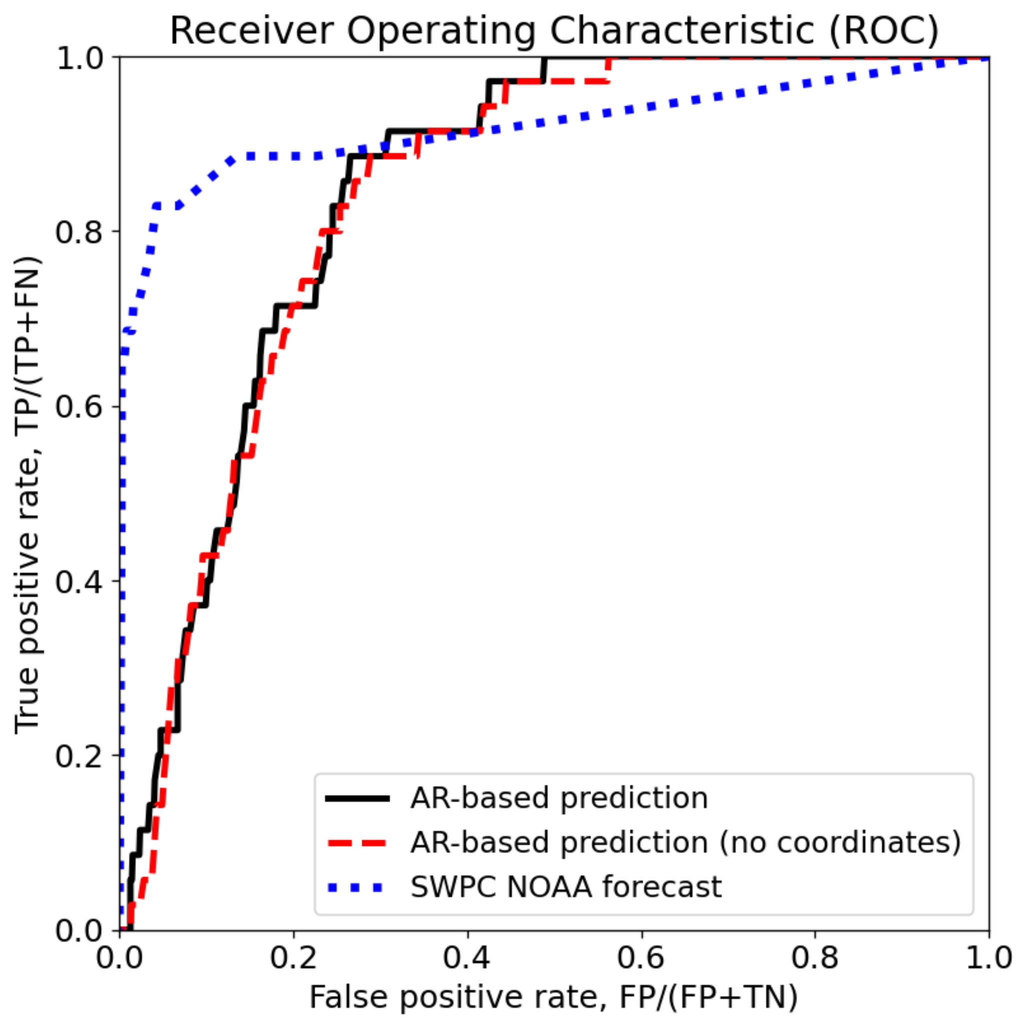
Understanding: ML prediction based on SHARP AR parameters only

	Cross-entropy loss	TSS	Best testing example			
			TP	TN	FP	FN
SHARP characteristics	0.477±0.003	0.574±0.006	32	787	391	3
Coordinates excluded	0.480±0.001	0.553±0.003	32	758	420	3
No behind-the-limb extension	0.589±0.018	0.473±0.013	30	742	436	5

- The neural network learns almost nothing if no behind-the-limb extension of active regions is implemented.
- The benefits from including AR coordinates are doubtful. Let’s look at ROC curves for these forecasts.

ROC curves (AR-based predictions)

- Inclusion of AR coordinates does not improve the prediction (although it has a higher TSS score).
- AR-based predictions are worse than the SWPC NOAA operational forecasts in the region of low false positive rates.
- However, there are certain advantages of AR-based forecasts: it is possible to predict all events and have the false positive rate never equal 1.

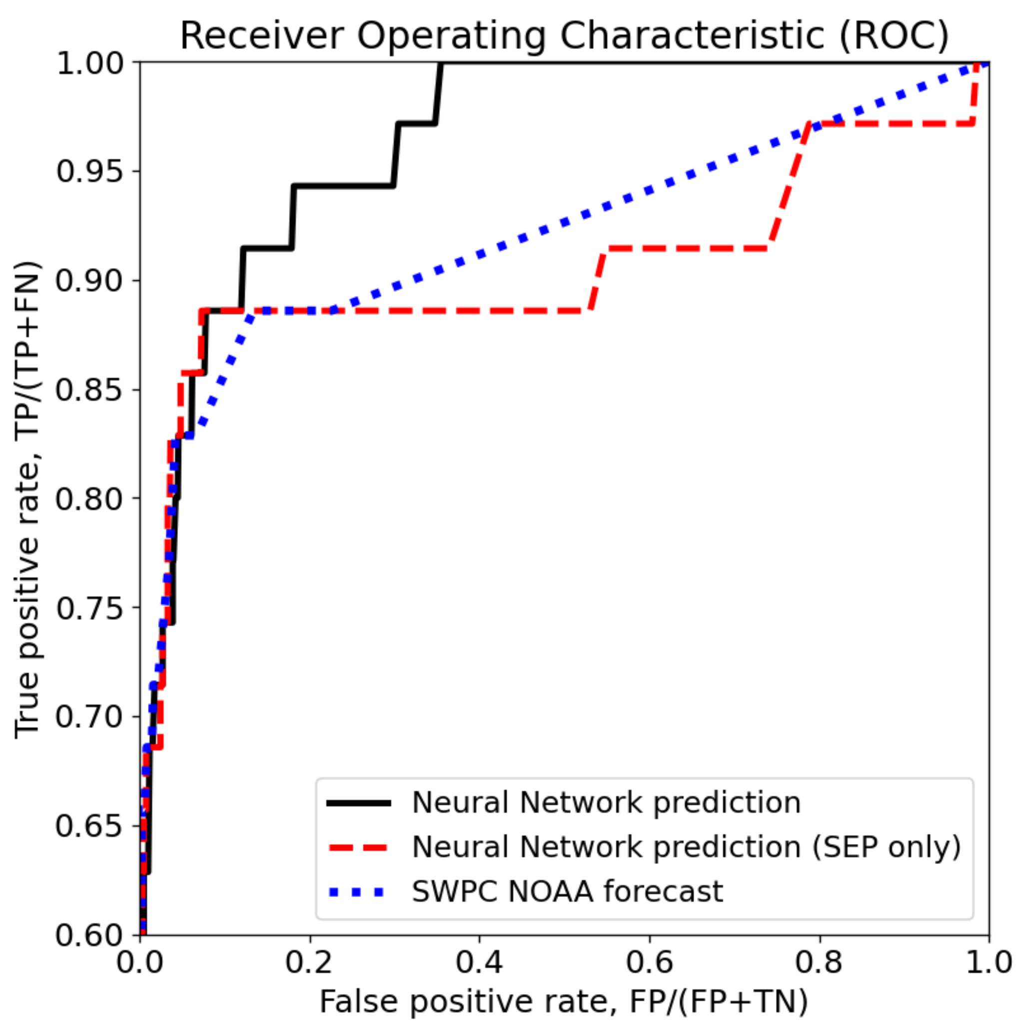


Understanding: inclusion/exclusion of features for ML prediction of SEPs

	Cross-entropy loss	TSS	Best testing example			
			TP	TN	FP	FN
All properties included	0.271±0.005	0.775±0.004	31	1054	124	4
AR information excluded	0.265±0.001	0.772±0.001	31	1046	132	4
SEP information excluded	0.497±0.008	0.499±0.007	31	734	444	4
SXR information excluded	0.282±0.015	0.765±0.015	31	1049	129	4
Radio burst information excluded	0.269±0.003	0.777±0.001	31	1054	124	4
SEP descriptors only	0.312±0.003	0.788±0.001	29	1131	47	6

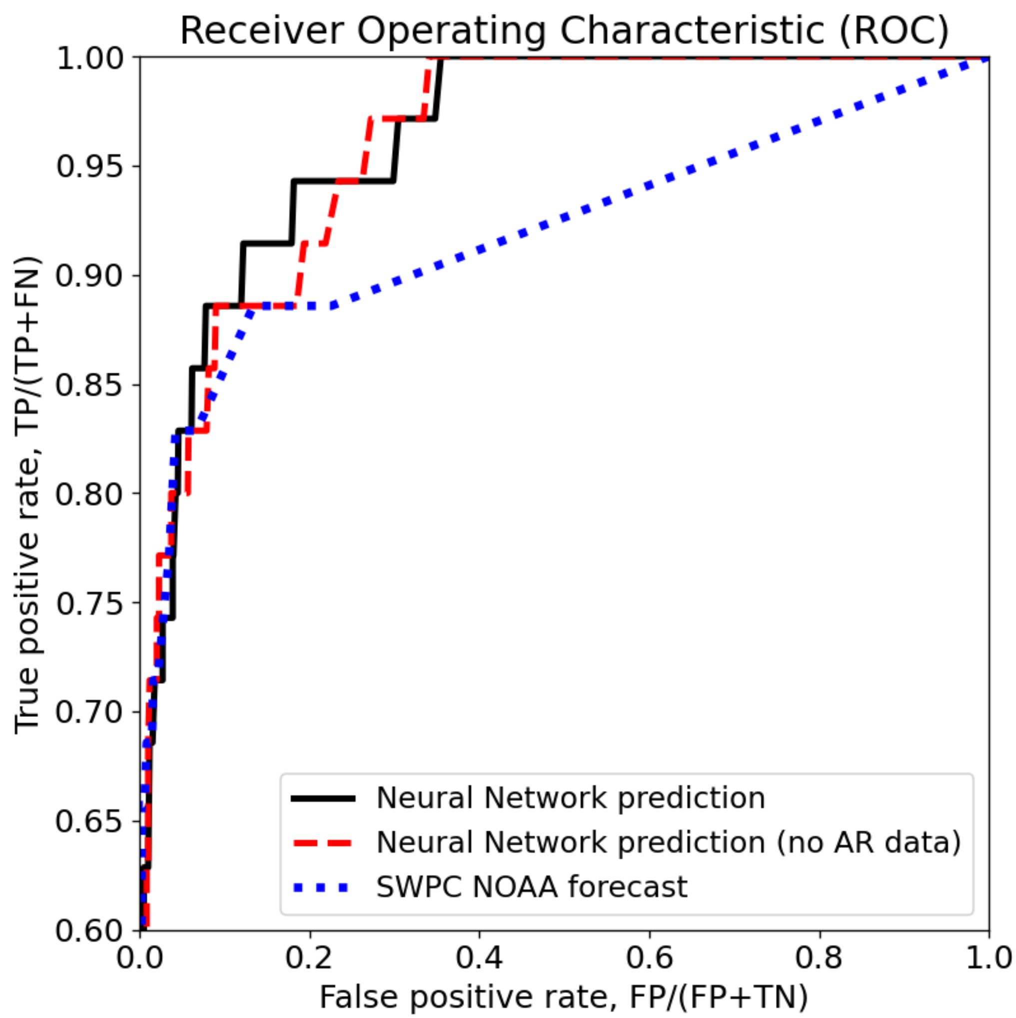
ROC curves (SEP characteristics)

- Inclusion of SEP characteristics is the most critical for network performance.
- The predictions behave very similarly to SWPC NOAA forecasts if trained on SEP characteristics only.



ROC curves (SEP characteristics)

- Exclusion of AR characteristics does not significantly affect the predictions
- There are two possible explanations:
  - All the necessary information is already contained in the SXR activity of the Sun.
  - Inclusion of AR dynamics is necessary for prediction capabilities.



Weighted TSS (WTSS) score

There is one more way to approach the “all-clear” forecast:

- True Skill Statistics score is defined as:

$$TSS = \frac{TP}{TP+FN} - \frac{FP}{FP+TN} = 1 - \frac{FN}{P} - \frac{FP}{N}$$

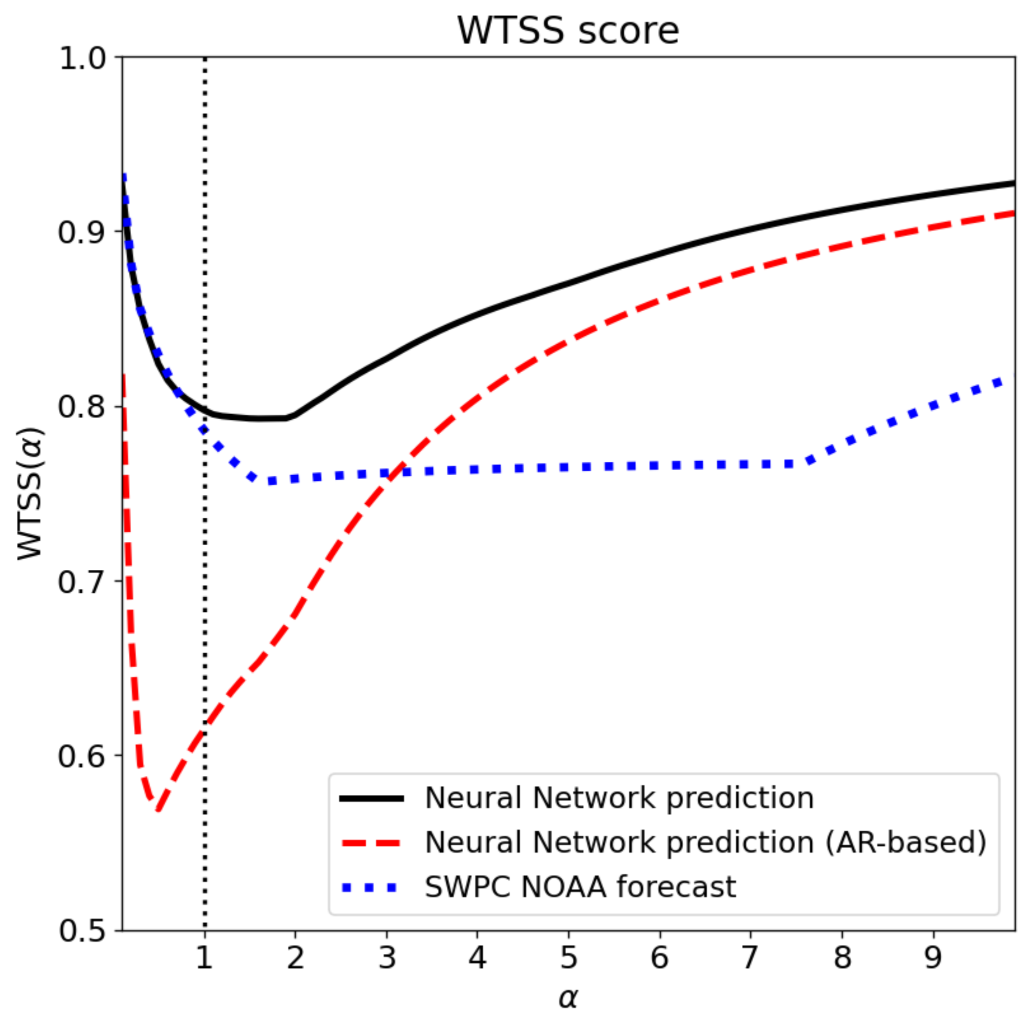
- Let’s apply weights to the missed event rate and the false alarm rate:

$$WTSS(\alpha) = 1 - \frac{2}{\alpha+1}\left(\alpha\frac{FN}{P} + \frac{FP}{N}\right)$$

- The parameter α indicates how much stronger our preference is for making the missed event rate lower with respect to the false alarm rate
- The WTSS(α) score has the same properties as the TSS score:
  - It ranges from -1 (totally wrong forecasts) to 1 (fully correct forecasts), where 0 corresponds to random guess forecasts.
  - It is not sensitive to the class-imbalance ratio.
  - $WTSS(1) = TSS$

WTSS results

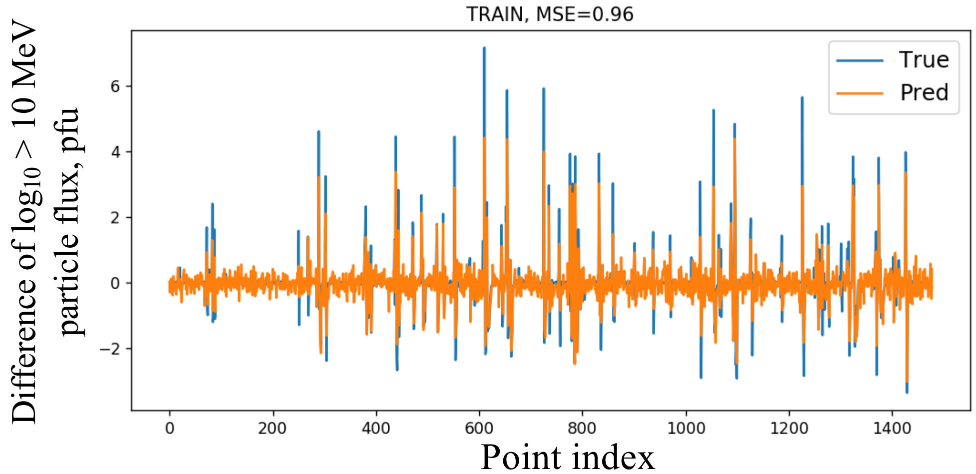
- The SWPC NOAA operational forecast and ML-based prediction are almost the same for α < 1.
- The ML-based prediction outperforms the SWPC NOAA operational forecast for α ≥ 1.
- Predictions based on AR parameters only have significantly lower scores than the other two predictions (SWPC NOAA and using all parameters) for small α.



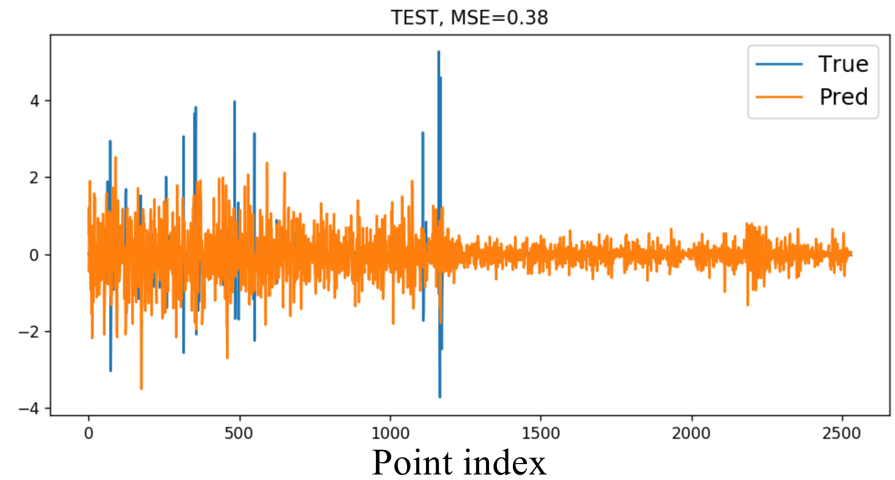
SEP flux prediction problem (E. Illarionov)

Goal: to predict <u>**a difference**</u> between the daily log-scaled flux values:

Performance on the train data set



Performance on the test data set



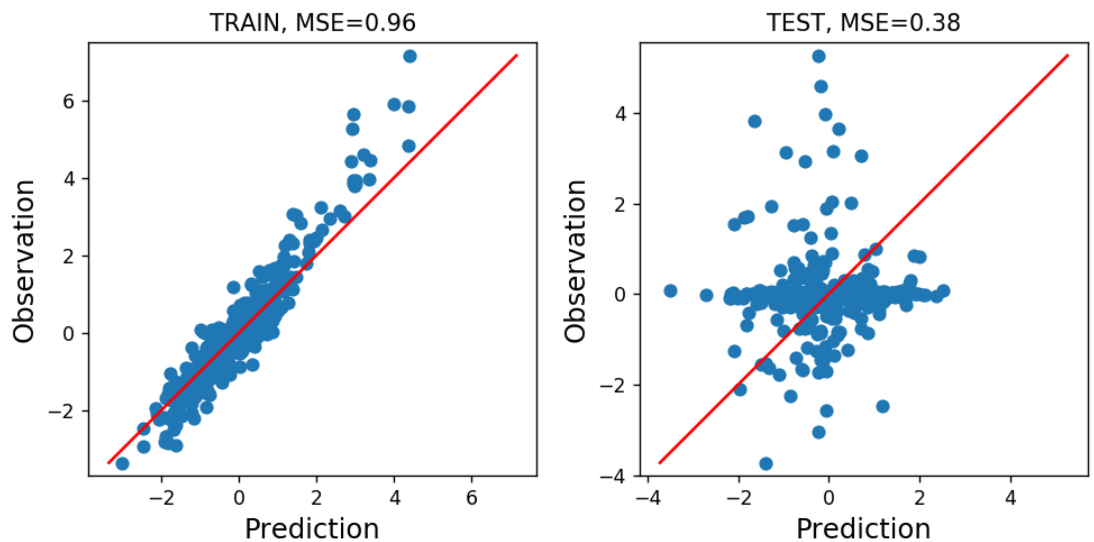
Courtesy: E. Illarionov

Notice that:

- If trained on the absolute values, the network tries to follow the persistence model (to predict the same flux as for the day before).
- Very common issue in time series forecasting

SEP flux prediction problem (E. Illarionov)

- Models tested: pure RNN-like models, 2D convolution models
- Cross-like pattern in the test data scatterplot shows that the model still behaves similarly to the persistent model.
- Ideas for the future:
  - Eliminate noisy features from the model (TAMs)
  - Add more data as a model input



Scatterplots for the difference of $\log_{10} > 10$ MeV particle flux, pfu. Courtesy: E. Illarionov.

Summary of the results

- Even a feature-based binary classification is an interesting problem!
- Inclusion of the western limb and far-side ARs is necessary if the AR features are considered in the forecast.
- Inclusion of SEP characteristics is the most critical for prediction.
- Exclusion of AR characteristics (in the form used in this study) does not seem to affect the predictions.
- Machine learning-based forecast seems to be very promising in situations when missed events are very undesirable ($\alpha > 1$ for WTSS). This is a good sign for "all-clear" forecast development!

Future development ideas

- Inclusion of more complex features of active regions and fluxes
- Inclusion and understanding the role of CME records in the prediction
- Inclusion of other features (temperatures and emission measures of flares, fluxes in other SEP energy channels, etc)
- Extension of the forecast to the Solar Cycle 23 (especially in case if characteristics of vector magnetic fields in ARs are not so important). Construction of the robust properly validated predictor.
- Exploration of other energy and particle flux thresholds, other timescales (shorter-term warnings are of particular interest), other targets (SPE peak flux and duration predictions)

Thank You for Your Attention!